Potential of CLARREO measurements for improving model ensemble based multi-decadal climate prediction: a statistical perspective

Yi Huang, Stephen Leroy, Richard Goody, and James Anderson

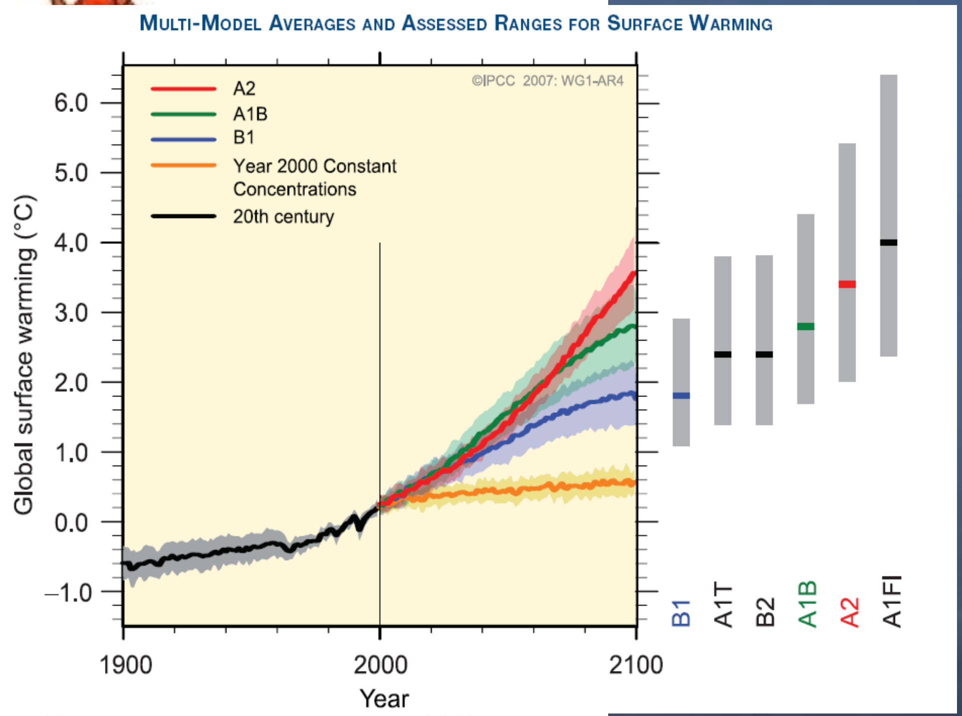
Harvard University

CLARREO workshop
May 18th, 2011

Climate prediction

- IPCC a
  - multip... ensem... project...
  - consi... uncerta... measu... inter- model

IPCC AR4

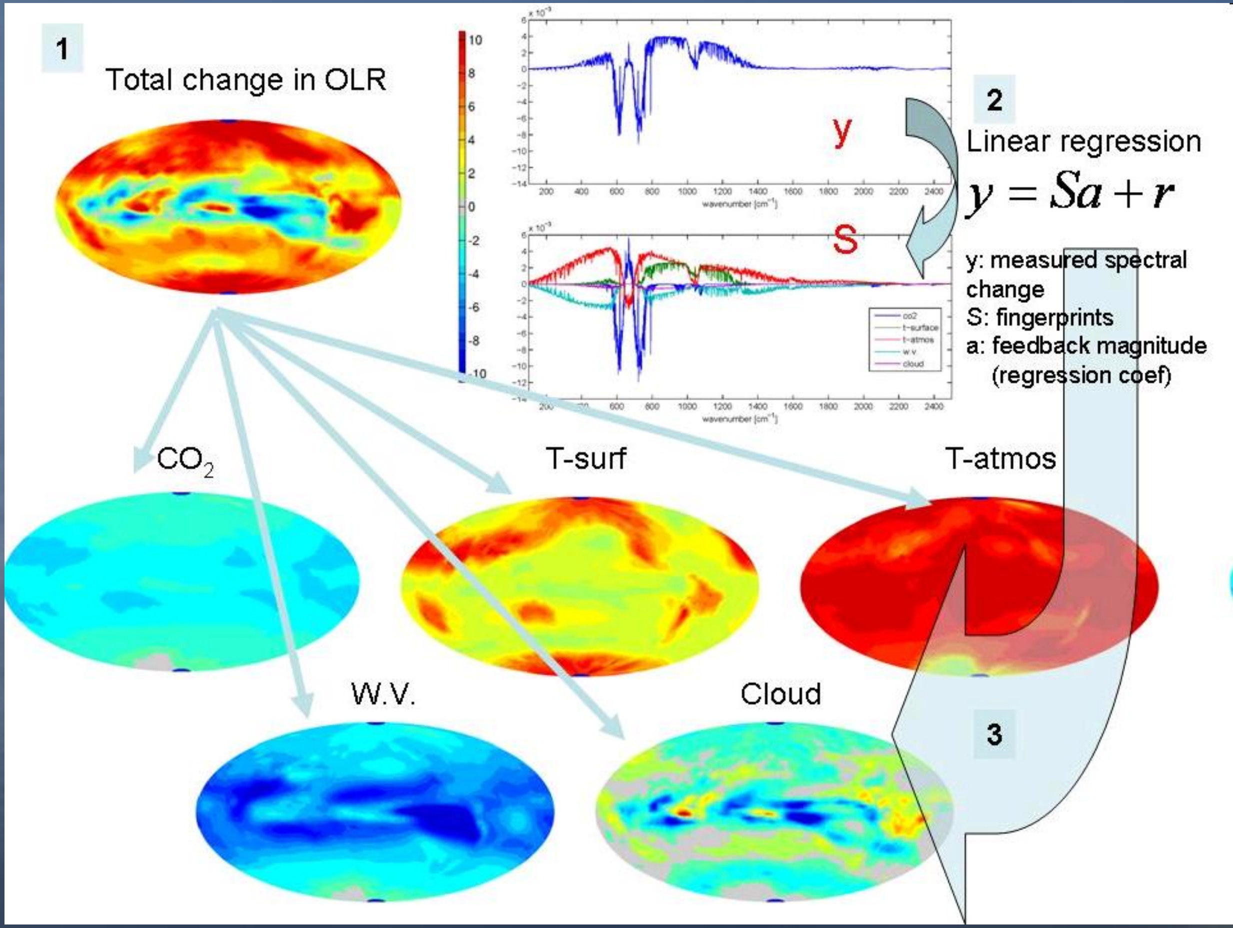


Climate prediction

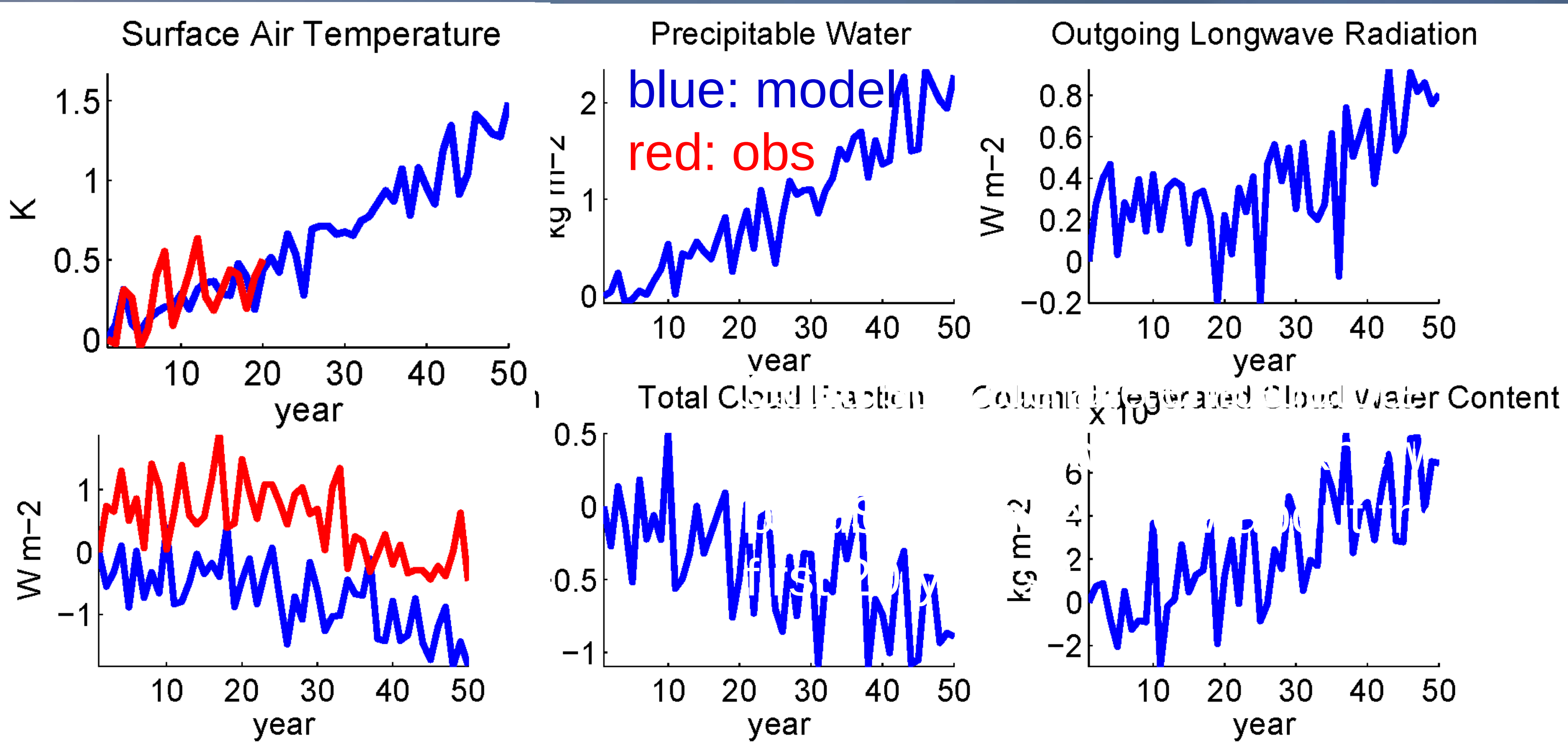
- How might CLARREO help?
  - Through testing and improving the models: Observation => model improvement => better prediction based on improved model
  - How to assess the impact of CLARREO NOW?
  - More generally, is it possible to improve the prediction through direct use of data? And what would be optimal data for this purpose?

  [Huang, Leroy and Goody, in press, PNAS]

Fingerprinting of the longwave climate feedbacks [Leroy et al., 2008; Huang et al. 2010]



- Model prediction – testable hypothesis
- Models predict various variables, a subset of which can be observed.
- Can a climate theory and its prediction of a variable of interest be improved by testing against ***available*** observation?



Bayesian Inference

$$P(A \mid B)P(B) = P(A, B) = P(B \mid A)P(A)$$

Probability of event A given event B

Joint probability of events A and B

Probability of event B given event A

A: Hypothesis

B: Data

Bayesian Inference

y: hypothetic change of interest

x: hypothetic change that can be measured

d: observation data - change actually measured

$$P(y \mid d) = \int P(x, y \mid d)\,dx \propto \int P(x, y)\,P(d \mid x)\,dx$$

Probability that the prediction of y is true given data

Probability that we would have observed d if the prediction of x was true

Prerequisite: the relationship between x and y (here, given by the CMIP3 models)

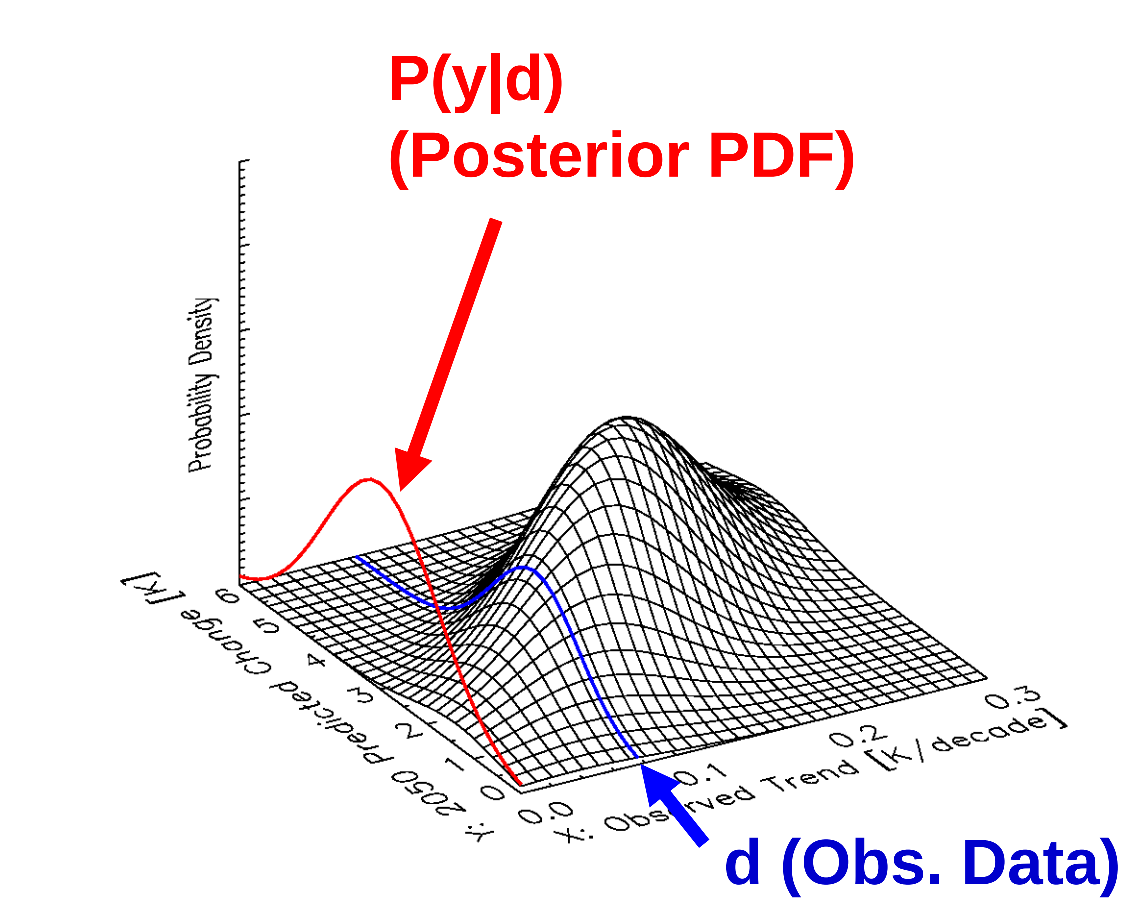
Bayesian Inference

y: hypothetic change of interest

x: hypothetic change that can be measured

d: observation data - change actually measured

$$P(y \mid d) = \int P(x, y \mid d) dx \propto \int P(x, y) P(d \mid x) dx$$



Bayesian Method

y: hypothetic change of interest

x: hypothetic change that can be measured

d: observation data - change actually measured

$$P(y \mid d) = \int P(x, y \mid d)dx \propto \int P(x, y)P(d \mid x)dx$$

Prior estimate

$$y \sim N(\mu_y, \sigma_y^2)$$

$$x \sim N(\mu_x, \sigma_x^2)$$

$$d \sim N(d, \sigma_d^2)$$

Measurement error

Prior uncertainty in the prediction
(model sensitivity difference + natural variability)

Posterior estimate

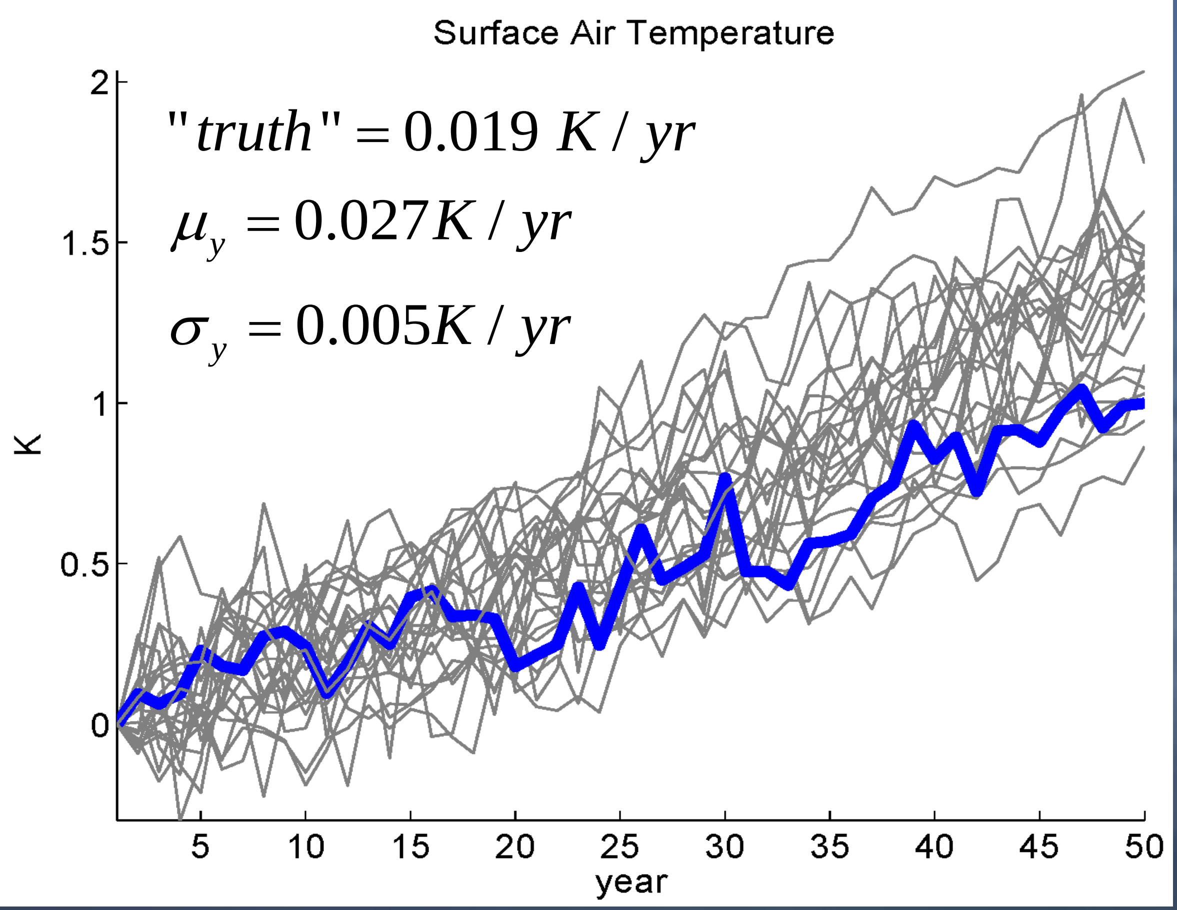
$$y \mid d \sim N(\mu_{y|d}, \sigma_{y|d}^2)$$

$$\mu_{y|d} = \mu_y + (d - \mu_x)\frac{\sigma_{xy}}{\sigma_x^2 + \sigma_d^2}$$

$$\sigma_{y|d}^2 = \sigma_y^2 (1 - \rho^2 \frac{\sigma_x^2}{\sigma_x^2 + \sigma_d^2})$$

Correlation between *x* and *y*

A “perfect model” test

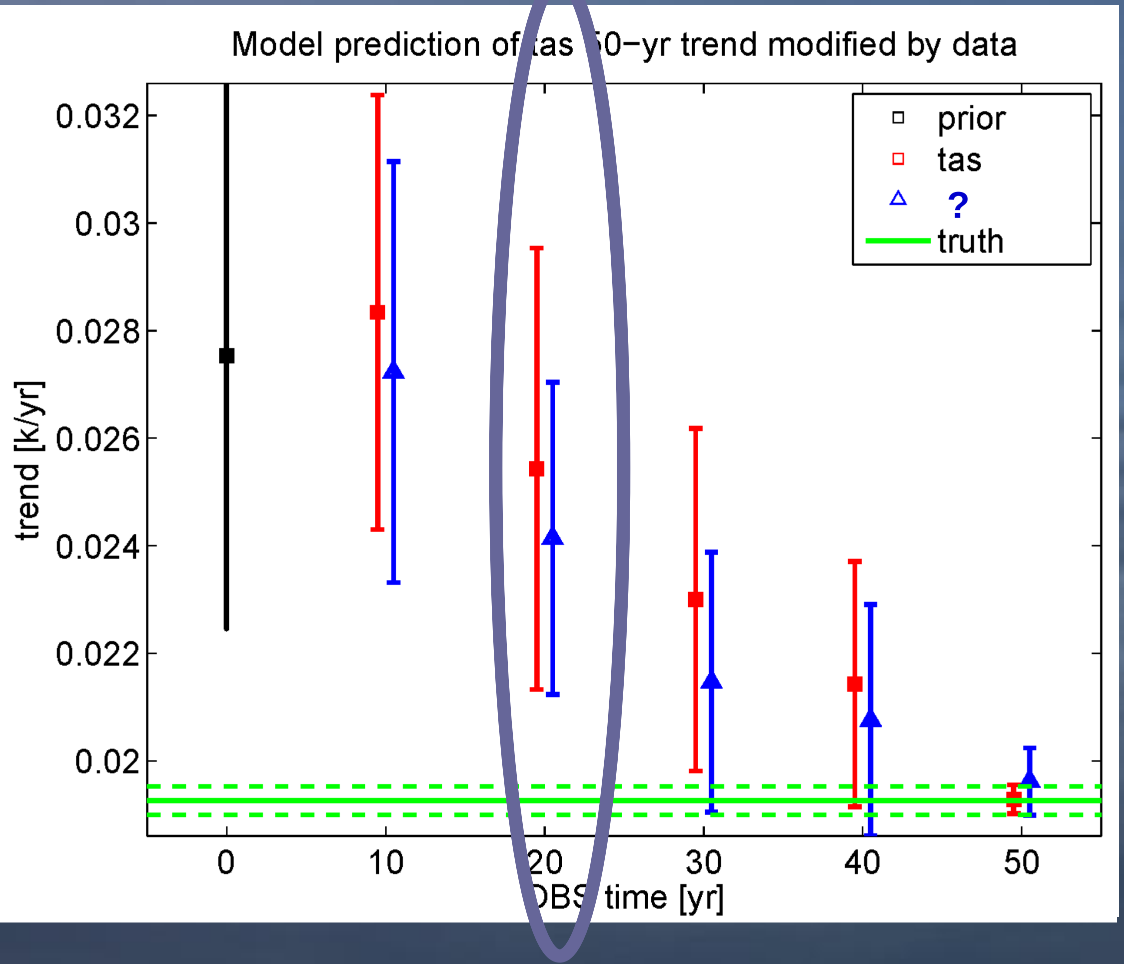


- CMIP3 (IPCC AR4) SresA1b experiment
  - One realization each model
  - x, y, d: all linear trends
- One model (ncar_pcm1) arbitrarily chosen to represent observational data, the “truth”.
- The prediction made by the rest models then validated against this “truth”.
- Target prediction: 50-year trend in the global mean surface air temperature

A "perfect model" test

$$\mu_{y|d} = \mu_y + (d - \mu_x)\frac{\sigma_{xy}}{\sigma_x^2 + \sigma_d^2}$$

$$\sigma_{y|d} = \sigma_y (1 - \rho^2 \frac{\sigma_x^2}{\sigma_x^2 + \sigma_d^2})^{1/2}$$



- When more and more data are obtained and used to modify the prior prediction according to the above equations, the posterior gets closer and closer to the truth, with less and less uncertainty.
- The best data type would provide the most constraint in the shortest observation time.

Calculations

- Target prediction (y): 50-year trend in surface air temperature
- Observation data (x): 20-year trends in:
  - (In situ) Surface air temperature (Tas), column integrated cloud water and ice, total cloud amount, precipitation, precipitable water (PW), surface downwelling (DLR) and upwelling longwave and shortwave radiation, and atmospheric temperature (Ta), relative humidity, specific humidity (q), and geopotential height (Z) at 500, 200 and 50 hPa levels, and
  - (satellite - CLARREO) Outgoing longwave radiation (OLR) and reflected shortwave (RS) radiation at TOA, (clear-sky) spectrally resolved OLR radiances, dry-pressures (P_{dry}) at 5.5, 10 and 14 Km.
- Metrics
  - Improvement in accuracy ($\Delta\mu$) and improvement in precision ($\Delta\sigma$):

$$\Delta\mu = \frac{\mu_y - \mu_{y|d}}{\mu_y - \mu_t} \qquad \Delta\sigma = \frac{\sigma_y - \sigma_{y|d}}{\sigma_y}$$

In situ

Data Type	ρ	Δσ	Δμ
Z_{500}	0.64	0.22	0.44
Tas	0.61	0.21	0.27
DLR	0.61	0.21	-0.47
PW	0.61	0.18	0.42
Z_{200}	0.56	0.17	0.39
Ta_{500}	0.53	0.15	0.44
q_{500}	0.45	0.10	0.43

Selection criterion: correlation significant at 95% confidence level.

$$\mu_{y|d} = \mu_y + (d - \mu_x)\frac{\sigma_{xy}}{\sigma_x^2 + \sigma_d^2}$$

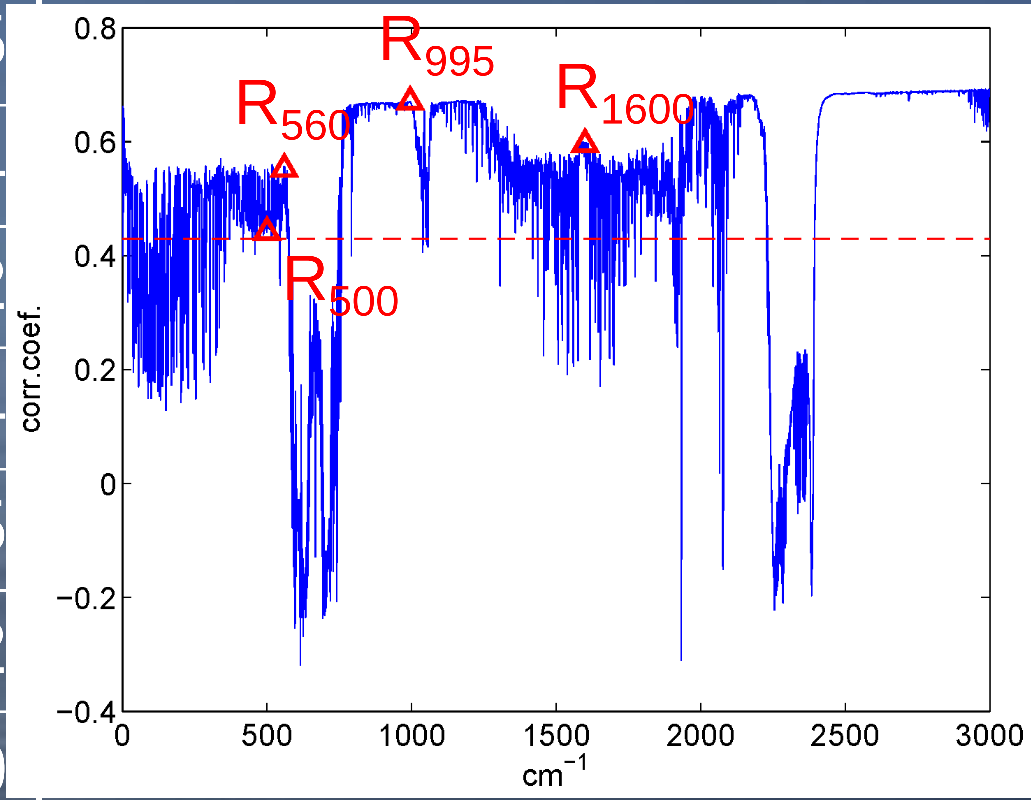
$$\sigma_{y|d} = \sigma_y (1 - \rho^2 \frac{\sigma_x^2}{\sigma_x^2 + \sigma_d^2})^{1/2}$$

$$\Delta\mu = \frac{\mu_y - \mu_{y|d}}{\mu_y - \mu_t}$$

$$\Delta\sigma = \frac{\sigma_y - \sigma_{y|d}}{\sigma_y}$$

Satellite

Data Type	ρ	$\Delta\sigma$	$\Delta\mu$
OLR	0.67	0.27	0.75
R_{995}	0.67	0.23	0.44
$P_{dry5.5}$	0.64	0.22	0.42
R_{1600}	0.60	0.18	0.34
R_{560}	0.55	0.15	0.35
RS	-0.47	0.13	0.32
R_{500}	0.44	0.09	0.30



Multiple data types used together to improve 50-year Tas trend prediction

Data Type (20-year observation)	Improvement in precision ($\Delta\sigma$)
Tas	0.21
All in situ data	0.35
LW Radiance data only	0.44
All satellite data	0.53

The results here indicate that CLARREO measurements are well chosen for providing powerful constraints on the precision of the ensemble prediction of surface temperature change.

Strict accuracy requirement on observing systems

$$\sigma_{y|d} = \sigma_y \left(1 - \rho^2 \frac{\sigma_x^2}{\sigma_x^2 + \sigma_d^2}\right)^{1/2}$$

20-yr trends	**σ_x** (inter-model difference)	**σ_d** (internal variability)	**σ_d** (WMO GCOS recommendations)
Tas [K / yr]	0.008	0.002	N/A
Ta [K / yr]	0.007	0.001	0.005
PW [kg m^{-2} / yr]	0.010	[illegible]	[illegible]
DLR [W m^{-2} / yr]	0.044	[illegible]	[illegible]
OLR [W m^{-2} / yr]	0.016	[illegible]	[illegible]
R_{995} [mW cm sr^{-1} m^{-2} / yr]	0.009	0.003	?CLARREO?
P_{dry} [hPa / yr]	0.009	0.002	?CLARREO?

σ ~ 0.033 K
=> 0.04 mW cm sr^{-1} m^{-2} radiance accuracy (at 995 cm^{-1} and 280K)
=> ~10 years is required to achieve the trend accuracy

Conclusions/discussions

- We present a methodology that provides constraints on the IPCC multi-model ensemble-based climate prediction by using observation data, and demonstrate that it can be used for selecting optimal data type for this purpose.
- 32 data types are examined for their potential for improving a 50-year surface air temperature trend prediction with data from earlier periods.
- Among the14 data types that are identified to be of significant potential,
  - The temperature data itself may not be the best data type for constraining surface air temperature prediction;
  - Most constraint is provided by OLR total flux and radiances.
- Given the sample size used to quantify $\rho(x,y)$, confidence on the ranking needs to be substantiated by a large ensemble.

Conclusions/discussions (cont')

- The results indicate that CLARREO measurements are well chosen for constraining ensemble prediction uncertainty and when used together may reduce the uncertainty in the 50-year temperature trend prediction by 50% in 20 years.
- Key to the improvement is the trend measurement accuracy, which constitutes a challenging requirement on most climate observing systems and is where the niche of CLARREO is.
- Yet to answer / improve:
  - Why is, e.g., ρ(OLR,Tas) high?
  - Gaussian assumption for P(x,y)
  - All-sky radiances
  - Optimal combination of data types (radiance selection)
  - Applied to real data

Backup

Bayesian Method

y: hypothetic change of interest

x: hypothetic change that can be measured

d: observation data - change actually measured

$$P(y \mid d) = \int P(x, y \mid d)dx \propto \int P(x, y)P(d \mid x)dx$$

Prior estimate

$$y \sim N(\mu_y, \sigma_y^2)$$

$$x \sim N(\mu_x, \sigma_x^2)$$

$$d \sim N(d, \sigma_d^2)$$

Measurement error

Prior uncertainty in the prediction
(model sensitivity difference + natural variability)

Posterior estimate

$$y \mid d \sim N(\mu_{y|d}, \sigma_{y|d}^2)$$

$$\mu_{y|d} = \mu_y + \Sigma_{yx}(\Sigma_{xx} + \Sigma_{dd})^{-1}(d - \mu_x)$$

$$\Sigma_{y|d} = \Sigma_{yy} - \Sigma_{yx}(\Sigma_{xx} + \Sigma_{dd})^{-1}\Sigma_{xy}$$

Correlation between *x* and *y*

CMIP3 720ppm stabilization (sresA1b) experiment

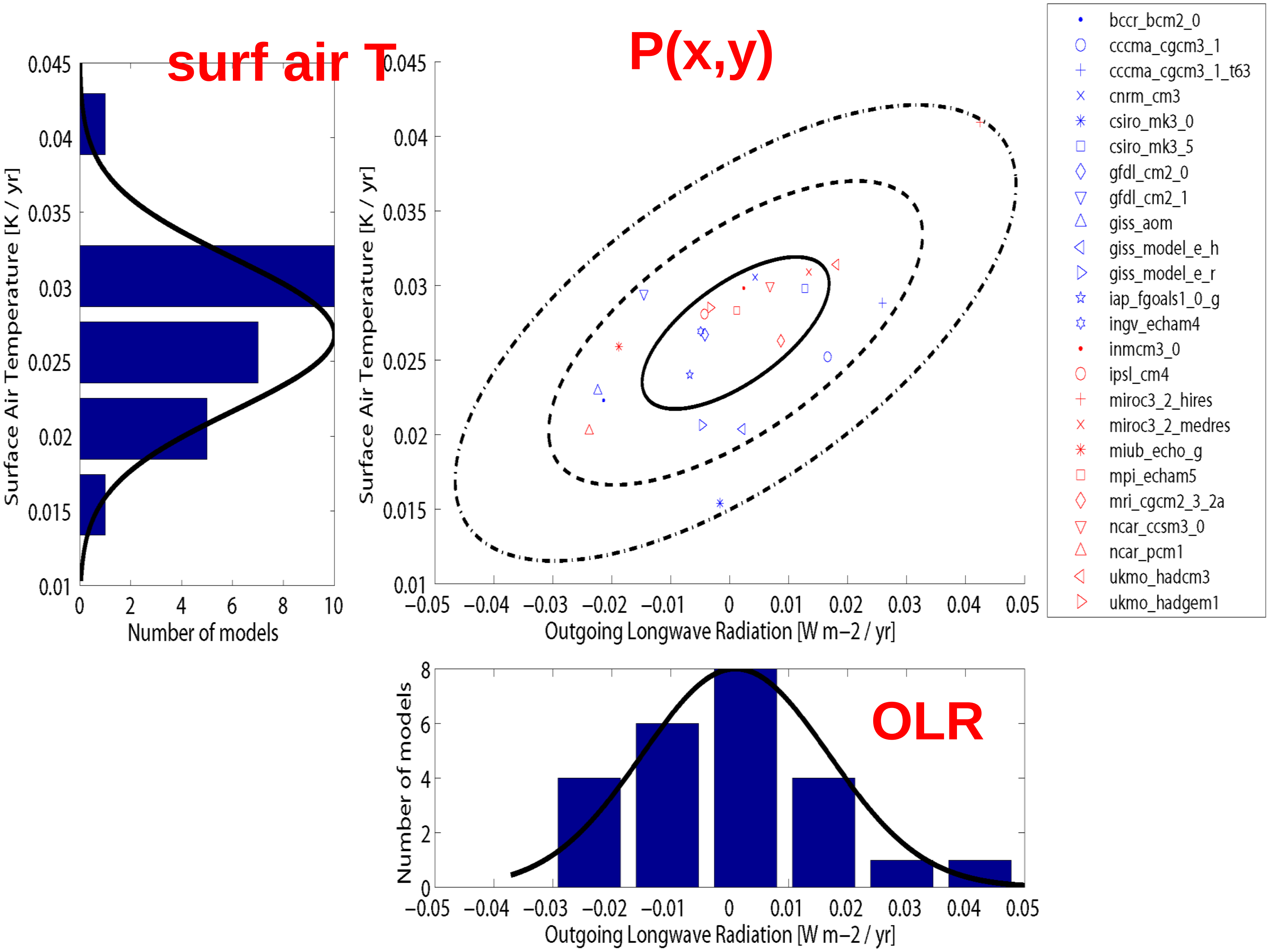
	model\variable	clivi	clwvi	clt	hur	hus	pr	prw	ps	psl	ta	tas	ts	rlds	rlus	rlut	rsds	rsus	rsdt	rsut	rtmt
NOR	bccr_bcm2_0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
CAN	cccma_cgcm3_1	5	5	5		5	5	5	5		5	5	5	5	5	5	5	5	5	5	5
	cccma_cgcm3_1_t63	1	1	1		1	1	1			1	1	1	1	1	1	1	1	1	1	1
FRA	cnrm_cm3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
AUS	csiro_mk3_0	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1
	csiro_mk3_5	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1
USA	gfdl_cm2_0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	gfdl_cm2_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	giss_aom			2		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
	giss_model_e_h	3	3	3	3	3	3	3	3	3	3	3	3			3	3	3	3	3	3
	giss_model_e_r	5	5	5	5	5	3	5	5	5	5	5	5	5	5	5	5	5	5	5	5
CHN	iap_fgoals1_0_g	3	3	3	3	3	3	3	1		3	3	3	3	3	3	3	3	3	3	3
ITA	ingv_echam4			1	1	1	1		1		1	1	1			1				1	1
RUS	inmcm3_0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FRA	ipsl_cm4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
JAP	miroc3_2_hires	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1
	miroc3_2_medres	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
GER	miub_echo_g		3	3			3	3				3	3	3	3	3	3	3	3	3	3
	mpi_echam5	4	4	4	4	4	4	4	3	3	4	4	3	3	4	4	4	4	4	4	4
JAP	mri_cgcm2_3_2a			1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
USA	ncar_ccsm3_0	7	7	7	7	7	7	7	6	5	7	7	7	7	6	7	7	7	7	7	7
	ncar_pcm1			4	4	4	4	4	1	1	4	4	4	4	4	2	4	4	4	2	2
UK	ukmo_hadcm3	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1
	ukmo_hadgem1	1	1	1	1	1	1		1		1	1	1	1	1	1	1	1	1	1	1

surface air temp.

- Source: WCRP CMIP3 multi-model data (ftp-esg.ucllnl.org)
- 24 GCMs
  - –Different number of runs from each GCM
  - –Some diagnostic variables are NOT available from some models
  - –Apparently wrong values in a few cases

Climate sensitivity (50-yr tas trend) of CMIP3 models

model	trend	rank	d_trend	rank	int.var.	rank
bccr_bcm2_0	0.02230	5	0.00385	15	0.10747	10
cccma_cgcm3_1	0.02531	8	0.00201	3	0.07267	5
cccma_cgcm3_1_t63	0.02883	16	0.00189	2	0.07255	4
cnrm_cm3	0.03056	21	0.00445	19	0.18647	22
csiro_mk3_0	0.01540	1	0.00294	8	0.11590	12
csiro_mk3_5	0.02980	19	0.00381	14	0.15004	20
gfdl_cm2_0	0.02670	11	0.00256	5	0.12322	16
gfdl_cm2_1	0.02944	18	0.00531	21	0.16838	21
giss_aom	0.02338	6	0.00211	4	0.05419	2
giss_model_e_h	0.01976	3	0.00267	6	0.06573	3
giss_model_e_r	0.02029	4	0.00149	1	0.04634	1
iap_fgoals1_0_g	0.02548	9	0.00392	16	0.19605	23
ingv_echam4	0.02693	12	0.00309	9	0.09537	8
inmcm3_0	0.02982	20	0.00336	12	0.11641	13
ipsl_cm4	0.02809	14	0.00281	7	0.12244	15
miroc3_2_hires	0.04093	24	0.00412	18	0.08123	7
miroc3_2_medres	0.03296	23	0.00557	23	0.12773	17
miub_echo_g	0.02516	7	0.00645	24	0.14153	19
mpi_echam5	0.02795	13	0.00542	22	0.20434	24
mri_cgcm2_3_2a	0.02644	10	0.00354	13	0.11938	14
ncar_ccsm3_0	0.02929	17	0.00320	11	0.07930	6
ncar_pcm1	0.01927	2	0.00319	10	0.10273	9
ukmo_hadcm3	0.03140	22	0.00399	17	0.12973	18
ukmo_hadgem1	0.02851	15	0.00459	20	0.11078	11



surf air T
P(x,y)
OLR
Surface Air Temperature [K / yr]
Number of models
Outgoing Longwave Radiation [W m−2 / yr]
Number of models
Outgoing Longwave Radiation [W m−2 / yr]
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0 2 4 6 8 10
−0.05 −0.04 −0.03 −0.02 −0.01 0 0.01 0.02 0.03 0.04 0.05
0 2 4 6 8
bccr_bcm2_0
cccma_cgcm3_1
cccma_cgcm3_1_t63
cnrm_cm3
csiro_mk3_0
csiro_mk3_5
gfdl_cm2_0
gfdl_cm2_1
giss_aom
giss_model_e_h
giss_model_e_r
iap_fgoals1_0_g
ingv_echam4
inmcm3_0
ipsl_cm4
miroc3_2_hires
miroc3_2_medres
miub_echo_g
mpi_echam5
mri_cgcm2_3_2a
ncar_ccsm3_0
ncar_pcm1
ukmo_hadcm3
ukmo_hadgem1